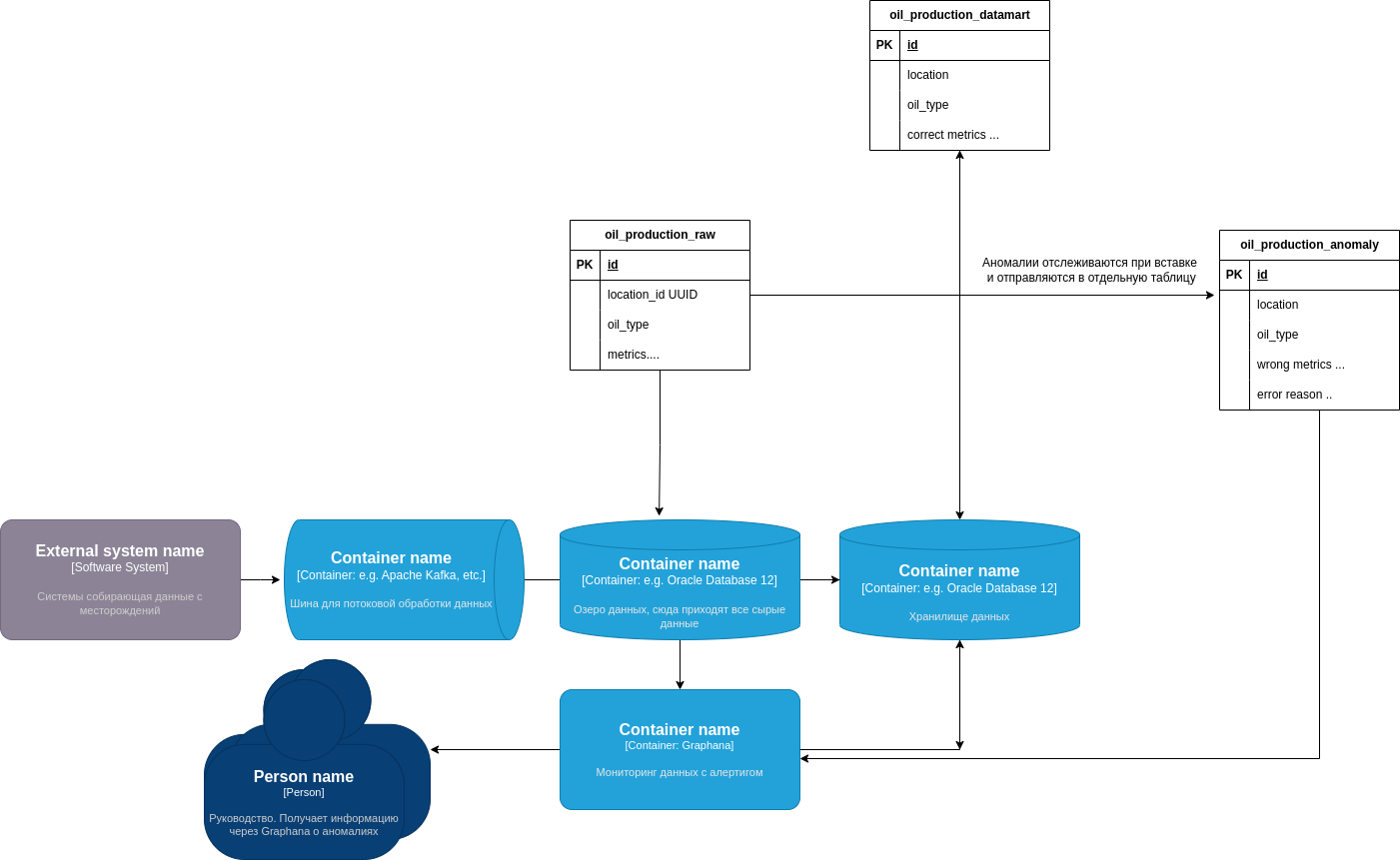
Лабораторная работа

Диаграмма приложения и кусок из бд для хранения выгружаемых данных и проверки данных

1. На рисунке выше приведена диаграмма сервисов и бд для хранения получаемых данных с месторождения нефти. Данные приходят ежеминутно из различных сторонних сервисов (или АСУТП) в шину, которая далее обрабатывает данные и передает в озеро данных для хранения. Далее данные через ETL процесс передаются в КХД, где прооисходит построение витрины данные по производству нефтепродуктов. На этом этапе происходит проверка качества данных: у таблицы в которую мы загружаем новые данные (oil\_production\_datamart) есть ограничения на некоторые поля и в случае конфликта неверные данные записываются в таблицу с аномалиями (oil\_production\_anomaly). Если для разных типов топлива разные предельные значения некоторых метрик, то эти метрики можно хранить в отдельной таблице и при джойне в витрину грузить только валидные данные, а не ошибочные грузить другим шагом ETL-процесса в oil\_production\_anomaly. Далее сервис мониторинга grafana смотрит на таблицу с аномалиями, и если в ней есть данные то она отправляет уведомления ответственным лицам для осведомления (в тг или почту).

2. Хранение пороговых значений описано выше. По сервису мониторинга выбор пал на grafana потому что это open-source с мощным коммунити, легко настроить мониторинг данные и железа, легко настроить нотификацию. Также есть встроенная поддержка широкого спектра баз данных. И широкий набор плагинов для кастомизации уведомлений, настройки дашбордов и прочее.